DISTRIBUTED COVERAGE GAMES FOR MOBILE VISUAL SENSOR NETWORKS

MINGHUI ZHU AND SONIA MARTíNEZ*

Abstract. Motivated by current challenges in data-intensive sensor networks, we formulate a coverage optimization problem for mobile visual sensors as a (constrained) repeated multi-player game. Each visual sensor tries to optimize its own coverage while minimizing the processing cost. We present two distributed learning algorithms where each sensor only remembers its own utility values and actions played during the last plays. These algorithms are proven to be convergent in probability to the set of (constrained) Nash equilibria and global optima of certain coverage performance metric, respectively.

1. Introduction. There is a widespread belief that continuous and pervasive monitoring will be possible in the near future with large numbers of networked, mobile, and wireless sensors. Thus, we are witnessing an intense research activity that focuses on the design of efficient control mechanisms for these systems. In particular, decentralized algorithms would allow sensor networks to react autonomously to changes in the environment with minimal human supervision.

A substantial body of research on sensor networks has concentrated on simple sensors that can collect scalar data; e.g., temperature, humidity or pressure data. Here, a main objective is the design of algorithms that can lead to optimal collective sensing through efficient motion control and communication schemes. However, scalar measurements can be insufficient in many situations; e.g., in automated surveillance or traffic monitoring applications. In contrast, data-intensive sensors such as cameras can collect visual data that are rich in information, thus having tremendous potential for monitoring applications, but at the cost of a higher processing overhead.

Precisely, this paper aims to solve a coverage optimization problem taking into account part of the sensing/processing trade-off. Coverage optimization problems have mainly been formulated as cooperative problems where each sensor benefits from sensing the environment as a member of a group. However, sensing may also require expenditure; e.g., the energy consumed or the time spent by image processing algorithms in visual networks. Because of this, we endow each sensor with a utility function that quantifies this trade-off, formulating a coverage problem as a variation of congestion games in [23].

Literature review. In broad terms, the problem studied here is related to a bevy of sensor location and planning problems in the Computational Geometry, Geometric Optimization, and Robotics literature. For example, different variations on the (combinatorial) Art Gallery problem include [22][25][27]. The objective here is how to find the optimum number of guards in a non-convex environment so that each point is visible from at least one guard. A related set of references for the deployment of mobile robots with omnidirectional cameras includes [10][9]. Unlike the Art Gallery classic algorithms, the latter papers assume that robots have local knowledge of the environment and no recollection of the past. Other related references on robot deployment in convex environments include [5][14] for anisotropic and circular footprints.

The paper [1] is an excellent survey on multimedia sensor networks where the state of the art in algorithms, protocols, and hardware is surveyed, and open research issues

*The authors are with Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr, La Jolla CA, 92093, {mizhu,soniamd}@ucsd.edu

are discussed in detail. As observed in [6], multimedia sensor networks enhance traditional surveillance systems by enlarging, enhancing, and enabling multi-resolution views. The investigation of coverage problems for static visual sensor networks is conducted in [4][11].

Another set of relevant references to this paper comprise those on the use of game-theoretic tools to (i) solve static target assignment problems, and (ii) devise efficient and secure algorithms for communication networks. In [16], the authors present a game-theoretic analysis of a coverage optimization problem for static sensor networks. This problem is equivalent to the weapon-target assignment problem in [21] which is nondeterministic polynomial-time complete. In general, the solution to assignment problems is hard from a combinatorial optimization viewpoint.

Game Theory and Learning in Games are used to analyze a variety of fundamental problems in; e.g., wireless communication networks and the Internet. An incomplete list of references includes [2] on power control, [24] on routing, and [26] on flow control. However, there has been limited research on how to employ Learning in Games to develop distributed algorithms for mobile sensor networks. One exception is the paper [15] where the authors establish a link between cooperative control problems (in particular, consensus problems), and games (in particular, potential games and weakly acyclic games).

Statement of contributions. The contributions of this paper pertain to both coverage optimization problems and Learning in Games. Compared with [13] and [14], this paper employs a more accurate sensing model and the results can be easily extended to include non-convex environments. Contrary to [13], we do not consider energy expenditure from sensor motions.

Regarding Learning in Games, we extend the use of the payoff-based learning dynamics in [17][18]. The coverage game we consider here is shown to be a (constrained) potential game. A number of learning rules; e.g., better (or best) reply dynamics and adaptive play, have been proposed to reach Nash equilibria in potential games. In these algorithms, each player must have access to the utility values induced by alternative actions. In our problem set-up; however, this information is unaccessible because of the information constraints caused by unknown rewards, motion and sensing limitations. To tackle this challenge, we develop two distributed payoff-based learning algorithms where each sensor only remembers its own utility values and actions played during the last plays.

In the first algorithm, at each time step, each sensor repeatedly updates its action synchronously, either trying some new action or selecting the action which corresponds to a higher utility value in the most recent two time steps. The first advantage of this algorithm over the payoff-based learning algorithms of [17][18] is its simpler dynamics, which reduces the computational complexity. Furthermore, the algorithm employs a diminishing exploration rate (in contrast to the constant one in [17][18]). The dynamically changing exploration rate renders the algorithm an inhomogeneous Markov chain (instead of the homogeneous ones in [17][18]). This technical change allows us to prove convergence in probability to the set of (constrained) Nash equilibria from which no agent is willing to unilaterally deviate. Thus, the property is substantially stronger than those in [17][18] where the algorithms are guaranteed to converge to Nash equilibria with a sufficiently large probability by choosing a sufficiently small exploration rate in advance.

The second algorithm is asynchronous. At each time step, only one sensor is active and updates its state by either trying some new action or selecting an action according to a Gibbs-like distribution from those played in last two time steps when it was active. The algorithm is shown to be convergent in probability to the set of global maxima of a coverage performance metric. Compared with the synchronous payoff-based log-linear learning algorithm in [17], this algorithm is asynchronous and simpler. Furthermore, rather than maximizing the associated potential function, the second algorithm optimizes a different global function which captures better a global trade-off between the overall network benefit from sensing and the total energy the network consumes. Again, by employing a diminishing exploration rate, our algorithm is guaranteed to have stronger convergence properties that the ones in [17].

2. Problem formulation. Here, we first review some basic game-theoretic concepts; see, for example [8]. This will allow us to formulate subsequently an optimal coverage problem for mobile visual sensor networks as a repeated multi-player game. We then introduce notation used throughout the paper.

2.1. Background in Game Theory. A strategic game $\Gamma := \langle V, A, U\rangle$ has three components:

1. A set V enumerating players $i \in V := \{1, \cdots, N\}$.
2. An action set $A := \prod_{i=1}^N A_i$ is the space of all actions vectors, where $s_i \in A_i$ is the action of player i and an (multi-player) action $s \in A$ has components $s_1, \ldots, s_N$.
3. The collection of utility functions U, where the utility function $u_i : A \to \mathbb{R}$ models player i's preferences over action profiles.

Denote by s_{-i} the action profile of all players other than i, and by $A_{-i} = \prod_{j\neq i} A_j$ the set of action profiles for all players except i. The concept of (pure) Nash equilibrium (NE, for short) is the most important one in Non-cooperative Game Theory [8] and is defined as follows.

Definition 2.1 (Nash equilibrium [8]). *Consider the strategic game Γ. An action profile $s^* := (s_i^*, s_{-i}^*)$ is a (pure) NE of the game Γ if $\forall i \in V$ and $\forall s_i \in A_i$, it holds that $u_i(s^*) \geq u_i(s_i, s_{-i}^*)$.*

An action profile corresponding to an NE represents a scenario where no player has incentive to unilaterally deviate. Potential Games form an important class of strategic games where the change in a player's utility caused by a unilateral deviation can be measured by a potential function.

Definition 2.2 (Potential game [20]). *The strategic game Γ is a potential game with potential function $\phi : A \to \mathbb{R}$ if for every $i \in V$, for every $s_{-i} \in A_{-i}$, and for every $s_i, s_i' \in A_i$, it holds that*

$$\phi(s_i, s_{-i}) - \phi(s_i', s_{-i}) = u_i(s_i, s_{-i}) - u_i(s_i', s_{-i}). \tag{2.1}$$

In conventional Non-cooperative Game Theory, all the actions in A_i always can be selected by player i in response to other players' actions. However, in the context of motion coordination, the actions available to player i will often be constrained to a state-dependent subset of A_i. In particular, we denote by $F_i(s_i, s_{-i}) \subseteq A_i$ the set of feasible actions of player i when the action profile is $s := (s_i, s_{-i})$. We assume that $F_i(s_i, s_{-i}) \neq \emptyset$. Denote $F(s) := \prod_{i\in V} F_i(s) \subseteq A$, $\forall s \in A$ and $F := \cup\{F(s) \mid s \in A\}$. The introduction of F leads naturally to the notion of constrained strategic game $\Gamma_{\text{res}} := \langle V, A, U, F\rangle$, and the following associated concepts.

Definition 2.3 (Constrained Nash equilibrium). *Consider the constrained strategic game Γ_{res}. An action profile s^* is a constrained (pure) NE of the game Γ_{res} if $\forall i \in V$ and $\forall s_i \in F_i(s_i^*, s_{-i}^*)$, it holds that $u_i(s^*) \geq u_i(s_i, s_{-i}^*)$.*

Definition 2.4 (Constrained potential game). *The game $\Gamma_{\rm res}$ is a constrained potential game with potential function $\phi(s)$ if for every $i \in V$, every $s_{-i} \in A_{-i}$, and every $s_i \in A_i$, the equality* (2.1) *holds for every $s_i' \in F_i(s_i, s_{-i})$.*

Observe that if s^* is an NE of the strategic game Γ, then it is also a constrained NE of the constrained strategic game $\Gamma_{\rm res}$. For any given strategic game, NE may not exist. However, the existence of NE in potential games is guaranteed [20]. Hence, any constrained potential game has at least one constrained NE.

2.2. Coverage problem formulation.

2.2.1. Mission space. We consider a convex 2-D mission space that is discretized into a (squared) lattice. We assume that each square of the lattice has unit dimensions. Each square will be labeled with the coordinate of its center $q = (q_x, q_y)$, where $q_x \in [q_{x_{\min}}, q_{x_{\max}}]$ and $q_y \in [q_{y_{\min}}, q_{y_{\max}}]$, for some integers $q_{x_{\min}},\ q_{y_{\min}}$, $q_{x_{\max}},\ q_{y_{\max}}$. Denote by $\mathcal{Q}$ the collection of all squares of the lattice.

We now define an associated location graph $\mathcal{G}_{\rm loc} := (\mathcal{Q}, E_{\rm loc})$ where $((q_x, q_y), (q_{x'}, q_{y'})) \in E_{\rm loc}$ if and only if $|q_x - q_{x'}| + |q_y - q_{y'}| = 1$ for $(q_x, q_y), (q_{x'}, q_{y'}) \in \mathcal{Q}$. Note that the graph $\mathcal{G}_{\rm loc}$ is undirected; i.e., $(q, q') \in E_{\rm loc}$ if and only if $(q', q) \in E_{\rm loc}$. The set of neighbors of q in $E_{\rm loc}$ is given by $\mathcal{N}_q^{\rm loc} := \{q' \in \mathcal{Q} \backslash \{q\} \mid (q, q') \in E_{\rm loc}\}$. We assume that the location graph $\mathcal{G}_{\rm loc}$ is fixed and connected, and denote its diameter by D.

Agents are deployed in $\mathcal{Q}$ to detect certain events of interest. As agents move in $\mathcal{Q}$ and process measurements, they will assign a numerical value $W_q \geq 0$ to the events in each square with center $q \in \mathcal{Q}$. If $W_q = 0$, then there is no significant event at the square with center q. The larger the value of W_q is, the more interest the set of events at the square with center q is of. Later, the amount W_q will be identified with a benefit of observing the point q. In this set-up, we assume the values W_q to be constant in time. Furthermore, W_q is not a prior knowledge to the agents, but the agents can measure this value through sensing the point q.

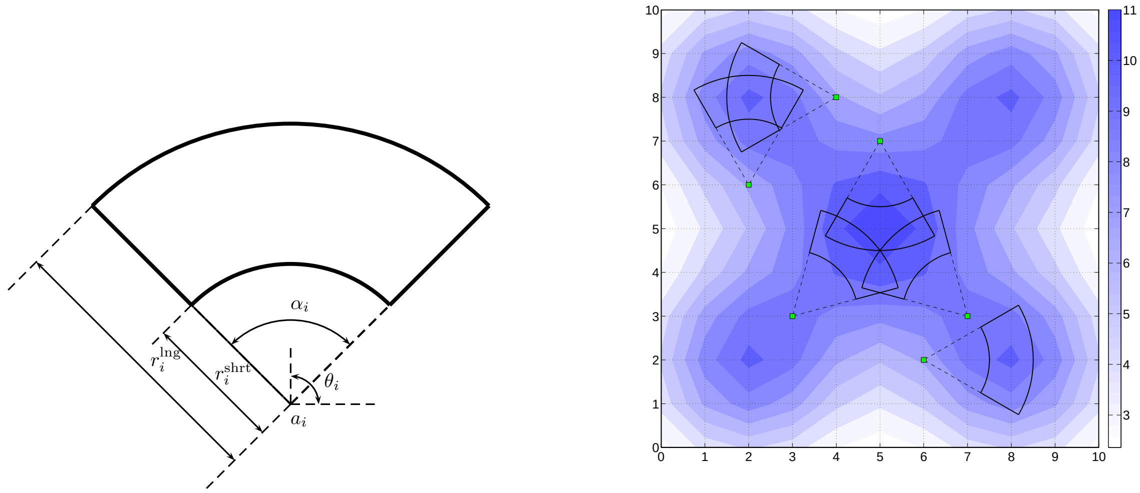
2.2.2. Modeling of the visual sensor nodes. Each mobile agent i is modeled as a point mass in $\mathcal{Q}$, with location $a_i := (x_i, y_i) \in \mathcal{Q}$. Each agent has mounted a pan-tilt-zoom camera, and can adjust its orientation and focal length.

The visual sensing range of a camera is directional, limited-range, and has a finite angle of view. Following a geometric simplification, we model the visual sensing region of agent i as an annulus sector in the 2-D plane; see Figure 2.1.

The visual sensor footprint is completely characterized by the following parameters: the position of agent i, $a_i \in \mathcal{Q}$, the camera orientation, $\theta_i \in [0, 2\pi)$, the camera angle of view, $\alpha_i \in [\alpha_{\min}, \alpha_{\max}]$, and the shortest range (resp. longest range) between agent i and the nearest (resp. farthest) object that can be recognized from the image, $r_i^{\rm shrt} \in [r_{\min}, r_{\max}]$ (resp. $r_i^{\rm lng} \in [r_{\min}, r_{\max}]$). The parameters $r_i^{\rm shrt}$, $r_i^{\rm lng}$, α_i can be tuned by changing the focal length ${\rm FL}_i$ of agent i's camera. In this way, $c_i := ({\rm FL}_i, \theta_i) \in [0, {\rm FL}_{\max}] \times [0, 2\pi)$ is the camera control vector of agent i. In what follows, we will assume that c_i takes values in a finite subset $\mathcal{C} \subset [0, {\rm FL}_{\max}] \times [0, 2\pi)$. An agent action is thus a vector $s_i := (a_i, c_i) \in \mathcal{A}_i := \mathcal{Q} \times \mathcal{C}$, and a multi-agent action is denoted by $s = (s_1, \ldots, s_N) \in \mathcal{A} := \Pi_{i=1}^N \mathcal{A}_i$.

Let $\mathcal{D}(a_i, c_i)$ be the visual sensor footprint of agent i. Now we can define a proximity sensing graph[1] $\mathcal{G}_{\rm sen}(s) := (V, E_{\rm sen}(s))$ as follows: the set of neighbors of agent i, $\mathcal{N}_i^{\rm sen}(s)$, is given as $\mathcal{N}_i^{\rm sen}(s) := \{j \in V \backslash \{i\} \mid \mathcal{D}(a_i, c_i) \cap \mathcal{D}(a_j, c_j) \cap \mathcal{Q} \neq \emptyset\}$.

[1] See [3] for a definition of proximity graph.



,

FIG. 2.1. *Visual sensor footprint and a configuration of the mobile sensor network*

Each agent is able to communicate with others to exchange information. We assume that the communication range of agents is $2r_{\max}$. This induces a $2r_{\max}$-disk communication graph $\mathcal{G}_{\rm comm}(s) := (V, E_{\rm comm}(s))$ as follows: the set of neighbors of agent i is given by $\mathcal{N}_i^{\rm comm}(s) := \{j \in V\backslash\{i\} \mid (x_i - x_j)^2 + (y_i - y_j)^2 \leq (2r_{\max})^2\}$. Note that $\mathcal{G}_{\rm comm}(s)$ is undirected and that $\mathcal{G}_{\rm sen}(s) \subseteq \mathcal{G}_{\rm comm}(s)$.

The motion of agents will be limited to a neighboring point in $\mathcal{G}_{\rm loc}$ at each time step. Thus, an agent feasible action set will be given by $\mathcal{F}_i(a_i) := (\{a_i\} \cup \mathcal{N}_{a_i}^{\rm loc}) \times \mathcal{C}$.

2.2.3. Coverage game. We now proceed to formulate a coverage optimization problem as a constrained strategic game. For each $q \in \mathcal{Q}$, we denote $n_q(s)$ as the cardinality of the set $\{k \in V \mid q \in \mathcal{D}(a_k, c_k) \cap \mathcal{Q}\}$; i.e., the number of agents which can observe the point q. The "profit" given by W_q will be equally shared by agents that can observe the point q. The benefit that agent i obtains through sensing is thus defined by $\sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)}$.

On the other hand, and as argued in [19], the processing of visual data can incur a higher cost than that of communication. This is in contrast with scalar sensor networks, where the communication cost dominates. With this observation, we model the energy consumption of agent i by $f_i(c_i) := \frac{1}{2}\alpha_i((r_i^{\rm lng})^2 - (r_i^{\rm shrt})^2)$. This measure corresponds to the area of the visual sensor footprint and can serve to approximate the energy consumption or the cost incurred by image processing algorithms.

We will endow each agent with a utility function that aims to capture the above sensing/processing trade-off. In this way, we define a utility function for agent i by

$$u_i(s) = \sum_{q \in \mathcal{D}(a_i, c_i) \cap \mathcal{Q}} \frac{W_q}{n_q(s)} - f_i(c_i).$$

Note that the utility function u_i is local over the visual sensing graph $\mathcal{G}_{\rm sen}(s)$; i.e., u_i is only dependent on the actions of $\{i\} \cup \mathcal{N}_i^{\rm sen}(s)$. With the set of utility functions $U_{\rm cov} = \{u_i\}_{i \in V}$, and feasible action set $\mathcal{F}_{\rm cov} = \Pi_{i=1}^N \bigcup_{a_i \in \mathcal{A}_i} \mathcal{F}_i(a_i)$, we now have all the ingredients to introduce the coverage game $\Gamma_{\rm cov} := \langle V, \mathcal{A}, U_{\rm cov}, \mathcal{F}_{\rm cov}\rangle$. This game is a variation of the congestion games introduced in [23].

LEMMA 2.5. *The coverage game $\Gamma_{\rm cov}$ is a constrained potential game with potential function*

$$\phi(s) = \sum_{q\in\mathcal{Q}} \sum_{\ell=1}^{n_q(s)} \frac{W_q}{\ell} - \sum_{i=1}^{N} f_i(c_i).$$

Proof. The proof is a slight variation of that in [23]. Consider any $s := (s_i, s_{-i}) \in \mathcal{A}$ where $s_i := (a_i, c_i)$. We fix $i \in V$ and pick any $s'_i = (a'_i, c'_i)$ from $\mathcal{F}_i(a_i)$. Denote $s' := (s'_i, s_{-i})$, $\Omega_1 := (\mathcal{D}(a_i, c_i)\backslash\mathcal{D}(a'_i, c'_i)) \cap \mathcal{Q}$ and $\Omega_2 := (\mathcal{D}(a'_i, c'_i)\backslash\mathcal{D}(a_i, c_i)) \cap \mathcal{Q}$. Observe that

$$\begin{aligned}
&\phi(s_i, s_{-i}) - \phi(s'_i, s_{-i})\\
&= \sum_{q\in\Omega_1}(\sum_{\ell=1}^{n_q(s)} \frac{W_q}{\ell} - \sum_{\ell=1}^{n_q(s')} \frac{W_q}{\ell}) + \sum_{q\in\Omega_2}(-\sum_{\ell=1}^{n_q(s)} \frac{W_q}{\ell} + \sum_{\ell=1}^{n_q(s')} \frac{W_q}{\ell}) - f_i(c_i) + f_i(c'_i)\\
&= \sum_{q\in\Omega_1} \frac{W_q}{n_q(s)} - \sum_{q\in\Omega_2} \frac{W_q}{n_q(s')} - f_i(c_i) + f_i(c'_i)\\
&= u_i(s_i, s_{-i}) - u_i(s'_i, s_{-i})
\end{aligned}$$

where in the second equality we utilize the fact that for each $q \in \Omega_1$, $n_q(s) = n_q(s')+1$, and each $q \in \Omega_2$, $n_q(s') = n_q(s) + 1$. ☐

We denote by $\mathcal{E}(\Gamma_{\rm cov})$ the set of constrained NEs of $\Gamma_{\rm cov}$. It is worth mentioning that $\mathcal{E}(\Gamma_{\rm cov}) \neq \emptyset$ due to the fact that $\Gamma_{\rm cov}$ is a constrained potential game.

Remark 2.1. *The assumptions of our problem formulation admit several extensions. For example, it is straightforward to extend our results to non-convex 3-D spaces. This is because the results that follow can also handle other shapes of the sensor footprint; e.g., a complete disk, a subset of the annulus sector. On the other hand, note that the coverage problem can be interpreted as a target assignment problem—here, the value $W_q \geq 0$ would be associated with the value of a target located at the point q.* •

2.3. Notations. In the following, we will use the Landau symbol, O, as in $O(\epsilon^k)$, for some $k \geq 0$. This implies that $0 < \lim_{\epsilon\to 0^+} \frac{O(\epsilon^k)}{\epsilon^k} < +\infty$. We denote by $\operatorname{diag}\mathcal{A} := \{(s, s) \in \mathcal{A}^2 \mid s \in \mathcal{A}\}$ and $\operatorname{diag}\mathcal{E}(\Gamma_{\rm cov}) := \{(s, s) \in \mathcal{A}^2 \mid s \in \mathcal{E}(\Gamma_{\rm cov})\}$.

Consider $a, a' \in \mathcal{Q}^N$ where $a_i \neq a'_i$ and $a_{-i} = a'_{-i}$ for some $i \in V$. The transition $a \to a'$ is feasible if and only if $(a_i, a'_i) \in E_{\rm loc}$. A feasible path from a to a' consisting of multiple feasible transitions is denoted by $a \Rightarrow a'$. Let $\diamond a := \{a' \in \mathcal{Q} \mid a \Rightarrow a'\}$ be the reachable set from a.

Let $s = (a, c), s' = (a', c') \in \mathcal{A}$ where $a_i \neq a'_i$ and $a_{-i} = a'_{-i}$ for some $i \in V$. The transition $s \to s'$ is feasible if and only if $s'_i \in \mathcal{F}_i(a)$. A feasible path from s to s' consisting of multiple feasible transitions is denoted by $s \Rightarrow s'$. Finally, $\diamond s := \{s' \in \mathcal{A} \mid s \Rightarrow s'\}$ will be the reachable set from s.

3. Distributed coverage learning algorithms and convergence results. In our coverage problem, we assume that W_q is unknown in advance. Furthermore, due to the limitations of motion and sensing, each agent is unable to obtain the information of W_q if the point q is outside its sensing range. These information constraints renders that each agent is unable to access to the utility values induced by alternative actions. Thus the action-based learning algorithms; e.g., better (or best) reply learning algorithm and adaptive play learning algorithm can not be employed to solve our coverage games. It motivates us to design distributed learning algorithms which only require the payoff received.

In this section, we come up with two distributed payoff-based learning algorithms, say *Distributed Inhomogeneous Synchronous Coverage Learning Algorithm* (DISCL, for short) and *Distributed Inhomogeneous Asynchronous Coverage Learning Algorithm* (DIACL, for short). We then present their convergence properties. Relevant algorithms include payoff-based learning algorithms proposed in [17][18].

3.1. Distributed Inhomogeneous Synchronous Coverage Learning Algorithm. For each $t \geq 1$ and $i \in V$, we define $\tau_i(t)$ as follows: $\tau_i(t) = t$ if $u_i(s(t)) \geq u_i(s(t-1))$, otherwise, $\tau_i(t) = t-1$. Here, $s_i(\tau_i(t))$ is the more successful action of agent i in last two steps. The main steps of the DISCL algorithm are the following:

Initialization: At $t = 0$, all agents are uniformly placed in $\mathcal{Q}$. Each agent i uniformly chooses its camera control vector c_i from the set $\mathcal{C}$, communicates with agents in $\mathcal{N}_i^{\rm sen}(s(0))$, and computes $u_i(s(0))$. At $t = 1$, all the agents keep their actions.

Update: At each time $t \geq 2$, each agent i updates its state according to the following rules:

- Agent i chooses the exploration rate $\epsilon(t) = t^{-\frac{1}{N(D+1)}}$ and compute $s_i(\tau_i(t))$.
- With probability $\epsilon(t)$, agent i experiments, and chooses the temporary action $s_i^{\rm tp} := (a_i^{\rm tp}, c_i^{\rm tp})$ uniformly from the set $\mathcal{F}_i(a_i(t)) \setminus \{s_i(\tau_i(t))\}$.
- With probability $1-\epsilon(t)$, agent i does not experiment, and sets $s_i^{\rm tp} = s_i(\tau_i(t))$.
- After $s_i^{\rm tp}$ is chosen, agent i moves to the position $a_i^{\rm tp}$ and sets the camera control vector to $c_i^{\rm tp}$.

Communication and computation: At position $a_i^{\rm tp}$, agent i communicates with agents in $\mathcal{N}_i^{\rm sen}(s_i^{\rm tp}, s_{-i}^{\rm tp})$, and computes $u_i(s_i^{\rm tp}, s_{-i}^{\rm tp})$ and $\mathcal{F}_i(a_i^{\rm tp})$.

4: Repeat Step 2 and 3.

Remark 3.1. *A variation of the* DISCL *algorithm corresponds to* $\epsilon(t) = \epsilon \in (0, \frac{1}{2}]$ *constant for all* $t \geq 2$. *If this is the case, we will refer to the algorithm as* Distributed Homogeneous Synchronous Coverage Learning Algorithm *(*DHSCL*, for short). Later, the convergence analysis of the* DISCL *algorithm will be based on the analysis of the* DHSCL *algorithm.* •

Denote the space $\mathcal{B} := \{(s, s') \in \mathcal{A} \times \mathcal{A} \mid s'_i \in \mathcal{F}_i(a_i),\ \forall i \in V\}$. Observe that $z(t) := (s(t-1), s(t))$ in the DISCL algorithm constitutes a time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ on the space $\mathcal{B}$. The following theorem states that the DISCL algorithm asymptotically converges to the set of $\mathcal{E}(\Gamma_{\rm cov})$ in probability.

THEOREM 3.1. *Consider the Markov chain* $\{\mathcal{P}_t\}$ *induced by the* DISCL *Algorithm. It holds that* $\lim_{t\to+\infty} \mathbb{P}(z(t) \in \operatorname{diag} \mathcal{E}(\Gamma_{\rm cov})) = 1$.

The proofs of Theorem 3.1 are provided in Section 4.

Remark 3.2. *The* DISCL *algorithm is simpler than the payoff-based learning algorithm proposed in [18], reducing the computational complexity. The algorithm studied in [18] converges the set of NEs with a arbitrarily high probability by choosing a arbitrarily small exploration rate* ϵ *in advance. However, it is difficult to derive the relation between the convergent probability and the exploration rate. It motivates us to utilize a diminishing exploration rate in the* DISCL *algorithm which induces a* *time-inhomogeneous Markov chain in contrast to a time-inhomogeneous Markov chain in [18]. This change renders a stronger convergence property, i.e., the convergence to the set of NEs in probability.* •

3.2. Distributed Inhomogeneous Asynchronous Coverage Learning Algorithm. Lemma 2.5 shows that the coverage game Γ_{cov} is a constrained potential game with potential function $\phi(s)$. However, this potential function is not a straightforward measure of the network coverage performance. On the other hand, the objective function $U_g(s) := \sum_{i\in V} u_i(s)$ captures the trade-off between the overall network benefit from sensing and the total energy the network consumes, and thus can be perceived as a more natural coverage performance metric. Denote by $S^* := \{s \mid \mathrm{argmax}_{s\in\mathcal{A}} U_g(s)\}$ as the set of global maximizers of $U_g(s)$. In this part, we present the DIACL algorithm which is convergent in probability to the set S^*.

Before that , we first introduce some notations for the DIACL algorithm. Denote by $\mathcal{B}'$ the space $\mathcal{B}' := \{(s, s') \in \mathcal{A}\times\mathcal{A} \mid s_{-i} = s'_{-i},\ s'_i \in \mathcal{F}_i(a_i) \text{ for some } i \in V\}$. For any $s^0, s^1 \in \mathcal{A}$ with $s^0_{-i} = s^1_{-i}$ for some $i \in V$, we denote

$$\Delta_i(s^1, s^0) := \frac{1}{2}\sum_{q\in\Omega_1}\frac{W_q}{n_q(s^1)} - \frac{1}{2}\sum_{q\in\Omega_2}\frac{W_q}{n_q(s^0)},$$

where $\Omega_1 := \mathcal{D}(a_i^1, c_i^1)\backslash\mathcal{D}(a_i^0, c_i^0)\cap\mathcal{Q}$ and $\Omega_2 := \mathcal{D}(a_i^0, c_i^0)\backslash\mathcal{D}(a_i^1, c_i^1)\cap\mathcal{Q}$, and

$$\begin{aligned}
&\rho_i(s^0, s^1) := u_i(s^1) - \Delta_i(s^1, s^0) - u_i(s^0) + \Delta_i(s^0, s^1),\\
&\Psi_i(s^0, s^1) := \max\{u_i(s^0) - \Delta_i(s^0, s^1), u_i(s^1) - \Delta_i(s^1, s^0)\},\\
&m^* := \max_{(s^0,s^1)\in\mathcal{B}, s_i^0\neq s_i^1}\{\Psi_i(s^0, s^1) - (u_i(s^0) - \Delta_i(s^0, s^1)), \frac{1}{2}\}.
\end{aligned}$$

It is easy to check that $\Delta_i(s^1, s^0) = -\Delta_i(s^0, s^1)$ and $\Psi_i(s^0, s^1) = \Psi_i(s^1, s^0)$. Assume that at each time instant, one of agents becomes active with equal probability. Denote by $\gamma_i(t)$ the last time instant before t when agent i was active. We then denote $\gamma_i^{(2)}(t) := \gamma_i(\gamma_i(t))$. The main steps of the DIACL algorithm are described in the following.

Initialization: At $t = 0$, all agents are uniformly placed in $\mathcal{Q}$. Each agent i uniformly chooses the camera control vector c_i from the set $\mathcal{C}$, and then communicates with agents in $\mathcal{N}_i^{\mathrm{sen}}(s(0))$ and computes $u_i(s(0))$. Furthermore, each agent i chooses $m_i \in (2m^*, Km^*]$ for some $K \geq 2$. At $t = 1$, all the sensors keep their actions.

Update: Assume that agent i is active at time $t \geq 2$. Then agent i updates its state according to the following rules:

- Agent i chooses the exploration rate $\epsilon(t) = t^{-\frac{1}{(D+1)(K+1)m^*}}$.
- With probability $\epsilon(t)^{m_i}$, agent i experiments and uniformly chooses $s_i^{\mathrm{tp}} := (a_i^{\mathrm{tp}}, c_i^{\mathrm{tp}})$ from the action set $\mathcal{F}_i(a_i(t)) \setminus \{s_i(t), s_i(\gamma_i^{(2)}(t)+1)\}$.
- With probability $1-\epsilon(t)^{m_i}$, agent i does not experiment and chooses s_i^{tp} according to the following probability distribution:

$$\mathbb{P}(s_i^{\mathrm{tp}} = s_i(t)) = \frac{1}{1+\epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}},$$

$$\mathbb{P}(s_i^{\mathrm{tp}} = s_i(\gamma_i^{(2)}(t)+1)) = \frac{\epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}}{1+\epsilon(t)^{\rho_i(s_i(\gamma_i^{(2)}(t)+1), s_i(t))}}.$$

• After $s_i^{\rm tp}$ is chosen, agent i moves to the position $a_i^{\rm tp}$ and sets its camera control vector to be $c_i^{\rm tp}$.

Communication and computation: At position $a_i^{\rm tp}$, the active agent i communicates with agents in $\mathcal{N}_i^{\rm sen}(s_i^{\rm tp}, s_{-i}(t))$, and computes $u_i(s_i^{\rm tp}, s_{-i}(t))$, $\Delta_i((s_i^{\rm tp}, s_{-i}(t)), s(\gamma_i(t)+1))$, $\mathcal{F}_i(a_i^{\rm tp})$.

4: Repeat Step 2 and 3.

Remark 3.3. *A variation of the* DIACL *algorithm corresponds to* $\epsilon(t) = \epsilon \in (0, \frac{1}{2}]$ *constant for all* $t \geq 2$. *If this is the case, we will refer to the algorithm as the* Distributed Homogeneous Asynchronous Coverage Learning Algorithm *(*DHACL*, for short). Later, we will base the convergence analysis of the* DIACL *algorithm on that of the* DHACL *algorithm.* •

Like the DISCL algorithm, $z(t) := (s(t-1), s(t))$ in the DIACL algorithm constitutes a time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ on the space $\mathcal{B}'$. The following theorem states that the convergence property of the DIACL algorithm.

THEOREM 3.2. *Consider the Markov chain* $\{\mathcal{P}_t\}$ *induced by the* DIACL *algorithm for the game* $\Gamma_{\rm cov}$. *Then it holds that* $\lim_{t\to+\infty} \mathbb{P}(z(t) \in \operatorname{diag} S^*) = 1$.

The proofs of Theorem 3.2 are provided in Section 4.

Remark 3.4. *The authors in [17] proposed a synchronous payoff-based log-linear learning algorithm. This algorithm is able to maximize the potential function of a potential game. While the* DIACL *algorithm is a variation of that in [17], and optimizes a different function* $U_g(s)$. *Furthermore, the convergence of the* DIACL *algorithm is in probability and stronger than the arbitrarily high probability [17] by choosing an arbitrarily small exploration rate in advance.* •

4. Convergence Analysis.

In this section, we prove Theorem 3.1 and 3.2 by appealing to the Theory of Resistance Trees in [28] and the results in strong ergodicity in [12]. Relevant papers include [17][18] where the Theory of Resistance Trees in [28] is first utilized to study the class of payoff-based learning algorithms, and [?][?][?] where the strong ergodicity theory is employed to characterize the convergence properties of time-inhomogeneous Markov chains.

4.1. Convergence analysis of the DISCL Algorithm.

We first utilize Theorem 6.6 to characterize the convergence properties of the associated DHSCL algorithm. This is essential for the analysis of the DISCL algorithm.

Observe that $z(t) := (s(t-1), s(t))$ in the DHSCL algorithm constitutes a time-homogeneous Markov chain $\{\mathcal{P}_t^\epsilon\}$ on the space $\mathcal{B}$. Consider $z, z' \in \mathcal{B}$. A feasible path from z to z' consisting of multiple feasible transitions of $\{\mathcal{P}_t^\epsilon\}$ is denoted by $z \Rightarrow z'$. The reachable set from z is denoted as $\diamond z := \{z' \in \mathcal{B} \mid z \Rightarrow z'\}$.

LEMMA 4.1. $\{\mathcal{P}_t^\epsilon\}$ *is a regular perturbation of* $\{\mathcal{P}_t^0\}$.

Proof. Consider a feasible transition $z^1 \to z^2$ with $z^1 := (s^0, s^1)$ and $z^2 := (s^1, s^2)$. Then we can define a partition of V as $\Lambda_1 := \{i \in V \mid s_i^2 = s_i^{\tau_i(0,1)}\}$ and $\Lambda_2 := \{i \in V \mid s_i^2 \in \mathcal{F}_i(a_i^1) \setminus \{s_i^{\tau_i(0,1)}\}\}$. The corresponding probability is given by

$$P_{z^1z^2}^\epsilon = \prod_{i\in\Lambda_1} (1-\epsilon) \times \prod_{j\in\Lambda_2} \frac{\epsilon}{|\mathcal{F}_i(a_i^1)| - 1}. \tag{4.1}$$

Hence, the resistance of the transition $z^1 \to z^2$ is $|\Lambda_2| \in \{0, 1, \cdots, N\}$ since

$$0 < \lim_{\epsilon\to0^+} \frac{P_{z^1z^2}^\epsilon}{\epsilon^{|\Lambda_2|}} = \prod_{j\in\Lambda_2} \frac{1}{|\mathcal{F}_i(a_i^1)| - 1} < +\infty.$$

We have that (A3) in Section 6.2 holds. It is not difficult to see that (A2) holds, and we are now in a position to verify (A1). Since $\mathcal{G}_{\rm loc}$ is undirected and connected, and multiple sensors can stay in the same position, then $\diamond a^0 = \mathcal{Q}^N$ for any $a^0 \in \mathcal{Q}$. Since sensor i can choose any camera control vector from $\mathcal{C}$ at each time, then $\diamond s^0 = \mathcal{A}$ for any $s^0 \in \mathcal{A}$. It implies that $\diamond z^0 = \mathcal{B}$ for any $z^0 \in \mathcal{B}$, and thus the Markov chain $\{\mathcal{P}_t^\epsilon\}$ is irreducible on the space $\mathcal{B}$.

It is easy to see that any state in $\mathrm{diag}\,\mathcal{A}$ has period 1. Pick any $(s^0, s^1) \in \mathcal{B}\backslash\mathrm{diag}\,\mathcal{A}$. Since $\mathcal{G}_{\rm loc}$ is undirected, then $s_i^0 \in \mathcal{F}_i(a_i^1)$ if and only if $s_i^1 \in \mathcal{F}_i(a_i^0)$. Hence, the following two paths are both feasible:

$$\begin{aligned}&(s^0, s^1) \rightarrow (s^1, s^0) \rightarrow (s^0, s^1)\\ &(s^0, s^1) \rightarrow (s^1, s^1) \rightarrow (s^1, s^0) \rightarrow (s^0, s^1).\end{aligned}$$

Hence, the period of the state (s^0, s^1) is 1. This proves aperiodicity of $\{\mathcal{P}_t^\epsilon\}$. Since $\{\mathcal{P}_t^\epsilon\}$ is irreducible and aperiodic, then (A1) holds. ☐

LEMMA 4.2. *For any $(s^0, s^0) \in \mathrm{diag}\,\mathcal{A} \setminus \mathrm{diag}\,\mathcal{E}(\Gamma_{\rm cov})$, there is a finite sequence of transitions from (s^0, s^0) to some $(s^*, s^*) \in \mathrm{diag}\,\mathcal{E}(\Gamma_{\rm cov})$ that satisfies*

$$\begin{aligned}\mathcal{L} := &(s^0, s^0) \overset{O(\epsilon)}{\rightarrow} (s^0, s^1) \overset{O(1)}{\rightarrow} (s^1, s^1) \overset{O(\epsilon)}{\rightarrow} (s^1, s^2)\\ &\overset{O(1)}{\rightarrow} (s^2, s^2) \overset{O(\epsilon)}{\rightarrow} \cdots \overset{O(\epsilon)}{\rightarrow} (s^{k-1}, s^k) \overset{O(1)}{\rightarrow} (s^k, s^k)\end{aligned}$$

where $(s^k, s^k) = (s^*, s^*)$ for some $k \geq 1$.

Proof. If $s^0 \notin \mathcal{E}(\Gamma_{\rm cov})$, there exists a sensor i with a action $s_i^1 \in \mathcal{F}_i(a_i^0)$ such that $u_i(s^1) > u_i(s^0)$ where $s_{-i}^0 = s_{-i}^1$. The transition $(s^0, s^0) \rightarrow (s^0, s^1)$ happens when only sensor i experiments, and its corresponding probability is $(1-\epsilon)^{N-1} \times \frac{\epsilon}{|\mathcal{F}_i(a_i^0)|-1}$. Since the function ϕ is the potential function of the game $\Gamma_{\rm cov}$, then we have that $\phi(s^1) - \phi(s^0) = u_i(s^1) - u_i(s^0)$ and thus $\phi(s^1) > \phi(s^0)$.

Since $u_i(s^1) > u_i(s^0)$ and $s_{-i}^0 = s_{-i}^1$, the transition $(s^0, s^1) \rightarrow (s^1, s^1)$ occurs when all sensors do not experiment, and the associated probability is $(1-\epsilon)^N$.

We repeat the above process and construct the path $\mathcal{L}$ with length $k \geq 1$. Since $\phi(s^i) > \phi(s^{i-1})$ for $i = \{1, \ldots, k\}$, then $s^i \neq s^j$ for $i \neq j$ and thus the path $\mathcal{L}$ has no loop. Since $\mathcal{A}$ is finite, then k is finite and thus $s^k = s^* \in \mathcal{E}(\Gamma_{\rm cov})$. ☐

A direct result of Lemma 4.1 is that for each ϵ, there exists a unique stationary distribution of $\{\mathcal{P}_t^\epsilon\}$, say $\mu(\epsilon)$. We now proceed to utilize Theorem 6.6 to characterize $\lim_{\epsilon\rightarrow 0^+} \mu(\epsilon)$.

PROPOSITION 4.3. *Consider the regular perturbation $\{\mathcal{P}_t^\epsilon\}$ of $\{\mathcal{P}_t^0\}$. Then $\lim\limits_{\epsilon\rightarrow 0^+} \mu(\epsilon)$ exists and the limiting distribution $\mu(0)$ is a stationary distribution of $\{\mathcal{P}_t^0\}$. Furthermore, the stochastically stable states (i.e., the support of $\mu(0)$) are contained in the set $\mathrm{diag}\,\mathcal{E}(\Gamma_{\rm cov})$.*

Proof. Notice that the stochastically stable states are contained in the recurrent communication classes of the unperturbed Markov chain that corresponds to the DHSCL Algorithm with $\epsilon = 0$. Thus the stochastically stable states are included in the set $\mathrm{diag}\,\mathcal{A} \subset \mathcal{B}$. Denote by $T_{\min}$ the minimum resistance tree and by h_v the root of $T_{\min}$. Each edge of $T_{\min}$ has resistance $0, 1, 2, \ldots$ corresponding to the transition probability $O(1), O(\epsilon), O(\epsilon^2), \ldots$. The state z' is the *successor* of the state z if and only if $(z, z') \in T_{\min}$. Like Theorem 3.2 in [18], our analysis will be slightly different from the presentation in 6.2. We will construct $T_{\min}$ over states in the set $\mathcal{B}$ (rather than $\mathrm{diag}\,\mathcal{A}$) with the restriction that all the edges leaving the states in $\mathcal{B}\backslash\mathrm{diag}\,\mathcal{A}$ have resistance 0. The stochastically stable states are not changed under this difference.

CLAIM 1. *For any* $(s^0, s^1) \in \mathcal{B} \setminus \operatorname{diag}\mathcal{A}$*, there is a finite path*

$$\mathcal{L}' := (s^0, s^1) \overset{O(1)}{\to} (s^1, s^2) \overset{O(1)}{\to} (s^2, s^2)$$

where $s_i^2 = s_i^{\tau_i(0,1)}$ *for all* $i \in V$*.*

Proof. These two transitions occur when all agents do not experiment. The corresponding probability of each transition is $(1-\epsilon)^N$. ∎

CLAIM 2. *The root* h_v *belongs to the set* $\operatorname{diag}\mathcal{A}$*.*

Proof. Suppose that $h_v = (s^0, s^1) \in \mathcal{B} \setminus \operatorname{diag}\mathcal{A}$. By Claim 1, there is a finite path $\mathcal{L}' := (s^0, s^1) \overset{O(1)}{\to} (s^1, s^2) \overset{O(1)}{\to} (s^2, s^2)$. We now construct a new tree T' by adding the edges of the path $\mathcal{L}'$ into the tree $T_{\min}$ and removing the redundant edges. The total resistance of adding edges is 0. Observe that the resistance of the removed edge exiting from (s^2, s^2) in the tree $T_{\min}$ is at least 1. Hence, the resistance of T' is strictly lower than that of $T_{\min}$, and we get to a contradiction.∎

CLAIM 3. *Pick any* $s^* \in \mathcal{E}(\Gamma_{\text{cov}})$ *and consider* $z := (s^*, s^*)$*,* $z' := (s^*, \tilde{s})$ *where* $\tilde{s} \neq s^*$*. If* $(z, z') \in T_{\min}$*, then the resistance of the edge* (z, z') *is some* $k \geq 2$*.*

Proof. Suppose the deviator in the transition $z \to z'$ is unique, say i. Then the corresponding transition probability is $O(\epsilon)$. Since $s^* \in \mathcal{E}(\Gamma_{\text{cov}})$ and $\tilde{s}_i \in \mathcal{F}_i(a_i^*)$, we have that $u_i(s_i^*, s_{-i}^*) \geq u_i(\tilde{s}_i, \tilde{s}_{-i})$, where $s_{-i}^* = \tilde{s}_{-i}$.

Since $z' \in \mathcal{B} \setminus \operatorname{diag}\mathcal{A}$, it follows from Claim 2 that the state z' can not be the root of $T_{\min}$ and thus has a successor z''. Note that all the edges leaving the states in $\mathcal{B} \setminus \operatorname{diag}\mathcal{A}$ have resistance 0. Then none experiments in the transition $z' \to z''$ and $z'' = (\tilde{s}, \hat{s})$ for some $\hat{s}$. Since $u_i(s_i^*, s_{-i}^*) \geq u_i(\tilde{s}_i, \tilde{s}_{-i})$ with $s_{-i}^* = \tilde{s}_{-i}$, we have $\hat{s} = s^*$ and thus $z'' = (\tilde{s}, s^*)$. Similarly, the state z'' must have a successor z''' and $z''' = z$. We then obtain a loop in $T_{\min}$ which contradicts that $T_{\min}$ is a tree.

It implies that at least two sensors experiment in the transition $z \to z'$. Thus the resistance of the edge (z, z') is at least 2.∎

CLAIM 4. *The root* h_v *belongs to the set* $\operatorname{diag}\mathcal{E}(\Gamma_{\text{cov}})$*.*

Proof. Suppose that $h_v = (s^0, s^0) \notin \operatorname{diag}\mathcal{E}(\Gamma_{\text{cov}})$. By Lemma 4.2, there is a finite path $\mathcal{L}$ connecting (s^0, s^0) and some $(s^*, s^*) \in \operatorname{diag}\mathcal{E}(\Gamma_{\text{cov}})$. We now construct a new tree T' by adding the edges of the path $\mathcal{L}$ into the tree $T_{\min}$ and removing the edges that leave the states in $\mathcal{L}$ in the tree $T_{\min}$. The total resistance of adding edges is k. Observe that the resistance of the removed edge exiting from (s^i, s^i) in the tree $T_{\min}$ is at least 1 for $i \in \{1, \cdots, k-1\}$. By Claim 3, the resistance of the removed edge leaving from (s^*, s^*) in the tree $T_{\min}$ is at least 2. The total resistance of removing edges is at least $k+1$. Hence, the resistance of T' is strictly lower than that of $T_{\min}$, and we get to a contradiction.∎

It follows from Claim 4 that the states in $\operatorname{diag}\mathcal{E}(\Gamma_{\text{cov}})$ have minimum stochastic potential. Since Lemma 4.1 shows that Markov chain $\{\mathcal{P}_t^\epsilon\}$ is a regularly perturbed Markov process, Proposition 4.3 is a direct result of Theorem 6.6. ∎

We are now ready to show Theorem 3.1.

Proof of Theorem 3.1:

CLAIM 5. *Condition (B2) in Theorem 6.5 holds.*

Proof. For each $t \geq 0$ and each $z \in X$, we defines the numbers

$$\sigma_z(\epsilon(t)) := \sum_{T \in G(z)} \prod_{(x,y) \in T} P_{xy}^{\epsilon(t)}, \quad \sigma_z^t = \sigma_z(\epsilon(t))$$

$$\mu_z(\epsilon(t)) := \frac{\sigma_z(\epsilon(t))}{\sum_{x \in X} \sigma_x(\epsilon(t))}, \quad \mu_z^t = \mu_z(\epsilon(t)).$$

Since $\{\mathcal{P}_t^\epsilon\}$ is a regular perturbation of $\{\mathcal{P}_t^0\}$, then it is irreducible and thus $\sigma_z^t > 0$. As Lemma 3.1 of Chapter 6 in [7], one can show that $(\mu^t)^T P^{\epsilon(t)} = (\mu^t)^T$. Therefore, condition (B2) in Theorem 6.5 holds.□

Claim 6. *Condition (B3) in Theorem 6.5 holds.*

Proof. We now proceed to verify condition (B3) in Theorem 6.5. To do that, let us first fix t, denote $\epsilon = \epsilon(t)$ and study the monotonicity of $\mu_z(\epsilon)$ with respect to ϵ. We write $\sigma_z(\epsilon)$ in the following form

$$\sigma_z(\epsilon) = \sum_{T\in G(z)} \prod_{(x,y)\in T} P_{xy}^\epsilon = \sum_{T\in G(z)} \prod_{(x,y)\in T} \frac{\alpha_{xy}(\epsilon)}{\beta_{xy}(\epsilon)} = \frac{\alpha_z(\epsilon)}{\beta_z(\epsilon)} \tag{4.2}$$

for some polynomials $\alpha_z(\epsilon)$ and $\beta_z(\epsilon)$ in ϵ. With (4.2) in hand, we have that $\sum_{x\in X}\sigma_x(\epsilon)$ and thus $\mu_z(\epsilon)$ are ratios of two polynomials in ϵ; i.e., $\mu_z(\epsilon) = \frac{\varphi_z(\epsilon)}{\beta(\epsilon)}$ where $\varphi_z(\epsilon)$ and $\beta(\epsilon)$ are polynomials in ϵ. The derivative of $\mu_z(\epsilon)$ is given by

$$\frac{\partial \mu_z(\epsilon)}{\partial \epsilon} = \frac{1}{\beta(\epsilon)^2}\Big(\frac{\partial \varphi_z(\epsilon)}{\partial \epsilon}\beta(\epsilon) - \varphi_z(\epsilon)\frac{\partial \beta(\epsilon)}{\partial \epsilon}\Big).$$

Note that the numerator $\frac{\partial\varphi_z(\epsilon)}{\partial\epsilon}\beta(\epsilon) - \varphi_z(\epsilon)\frac{\partial\beta(\epsilon)}{\partial\epsilon}$ is a polynomial in ϵ. Denote by $\iota_z \neq 0$ the coefficient of the leading term of $\frac{\partial\varphi_z(\epsilon)}{\partial\epsilon} - \varphi_z(\epsilon)\frac{\partial\beta(\epsilon)}{\epsilon}$. The leading term dominates $\frac{\partial\varphi_z(\epsilon)}{\partial\epsilon} - \varphi_z(\epsilon)\frac{\partial\beta(\epsilon)}{\epsilon}$ when ϵ is sufficiently small. Thus there exists $\epsilon_z > 0$ such that the sign of $\frac{\partial\mu_z(\epsilon)}{\partial\epsilon}$ is the sign of ι_z for all $0 < \epsilon \leq \epsilon_z$. Let $\epsilon^* = \max_{z\in X}\epsilon_z$.

Since $\epsilon(t)$ strictly decreases to zero, then there is a unique finite time instant t^* such that $\epsilon(t^*) = \epsilon^*$ (if $\epsilon(0) < \epsilon^*$, then $t^* = 0$). Since $\epsilon(t)$ is strictly decreasing, we can define a partition of X as follows:

$$\begin{aligned}
\Xi_1 &:= \{z \in X \mid \mu_z(\epsilon(t)) > \mu_z(\epsilon(t+1)), \quad \forall t \in [t^*, +\infty)\},\\
\Xi_2 &:= \{z \in X \mid \mu_z(\epsilon(t)) < \mu_z(\epsilon(t+1)), \quad \forall t \in [t^*, +\infty)\}.
\end{aligned}$$

We are now ready to verify (B3) of Theorem 6.5. Since $\{\mathcal{P}_t^\epsilon\}$ is a regular perturbed Markov chain of $\{\mathcal{P}_t^0\}$, it follows from Theorem 6.6 that $\lim_{t\to+\infty}\mu_z(\epsilon(t)) = \mu_z(0)$, and thus it holds that

$$\begin{aligned}
&\sum_{t=0}^{+\infty}\sum_{z\in X}\|\mu_z^t - \mu_z^{t+1}\| = \sum_{t=0}^{+\infty}\sum_{z\in X}|\mu_z(\epsilon(t)) - \mu_z(\epsilon(t+1))|\\
&= \sum_{t=0}^{t^*}\sum_{z\in X}|\mu_z(\epsilon(t)) - \mu_z(\epsilon(t+1))| + \sum_{t=t^*+1}^{+\infty}\Big(\sum_{z\in\Xi_1}\mu_z(\epsilon(t)) - \sum_{z\in\Xi_1}\mu_z(\epsilon(t+1))\Big)\\
&+ \sum_{t=t^*+1}^{+\infty}\Big(1 - \sum_{z\in\Xi_1}\mu_z(\epsilon(t+1)) - \big(1 - \sum_{z\in\Xi_1}\mu_z(\epsilon(t))\big)\Big)\\
&= \sum_{t=0}^{t^*}\sum_{z\in X}|\mu_z(\epsilon(t)) - \mu_z(\epsilon(t+1))| + 2\sum_{z\in\Xi_1}\mu_z(\epsilon(t^*+1)) - 2\sum_{z\in\Xi_1}\mu_z(0) < +\infty.
\end{aligned}$$

□

Claim 7. *Condition (B1) in Theorem 6.5 holds.*

Proof. Denote by $P^{\epsilon(t)}$ the transition matrix of $\{\mathcal{P}_t\}$. As in (4.1), the probability of the feasible transition $z^1 \to z^2$ is given by

$$P_{z^1z^2}^{\epsilon(t)} = \prod_{i\in\Lambda_1}(1-\epsilon(t)) \times \prod_{j\in\Lambda_2}\frac{\epsilon(t)}{|\mathcal{F}_i(a_i^1)| - 1}.$$

Observe that $|\mathcal{F}_i(a_i^1)| \leq 5|\mathcal{C}|$. Since $\epsilon(t)$ is strictly decreasing, there is $t_0 \geq 1$ such that t_0 is the first time when $1-\epsilon(t) \geq \frac{\epsilon(t)}{5|\mathcal{C}|-1}$. Then for all $t \geq t_0$, it holds that

$$P^{\epsilon(t)}_{z^1 z^2} \geq (\frac{\epsilon(t)}{5|\mathcal{C}|-1})^N.$$

Denote $P(m,n) := \prod_{t=m}^{n-1} P^{\epsilon(t)}$, $0 \leq m < n$. Pick any $z \in \mathcal{B}$ and let $u_z \in \mathcal{B}$ be such that $P_{u_z z}(t, t+D+1) = \min_{x\in\mathcal{B}} P_{xz}(t,t+D+1)$. Consequently, it follows that for all $t \geq t_0$,

$$\min_{x\in\mathcal{B}} P_{xz}(t,t+D+1) = \sum_{i_1\in\mathcal{B}} \cdots \sum_{i_D\in\in\mathcal{B}} P^{\epsilon(t)}_{u_z i_1} \cdots P^{\epsilon(t+D-1)}_{i_{D-1} i_D} P^{\epsilon(t+D)}_{i_D z}$$
$$\geq P^{\epsilon(t)}_{u_z i_1} \cdots P^{\epsilon(t+D-1)}_{i_{D-1} i_D} P^{\epsilon(t+D)}_{i_D z} \geq \prod_{i=0}^{D} (\frac{\epsilon(t+i)}{5|\mathcal{C}|-1})^N \geq (\frac{\epsilon(t)}{5|\mathcal{C}|-1})^{(D+1)N}$$

where in the last inequality we use $\epsilon(t)$ begin strictly decreasing. Then we have

$$1-\lambda(P(t,t+D+1)) = \min_{x,y\in\mathcal{B}} \sum_{z\in\mathcal{B}} \min\{P_{xz}(t,t+D+1), P_{yz}(t,t+D+1)\}$$
$$\geq \sum_{z\in\mathcal{B}} P_{u_z z}(t,t+D+1) \geq |\mathcal{B}|(\frac{\epsilon(t)}{5|\mathcal{C}|-1})^{(D+1)N}.$$

Choose $k_i := (D+1)i$ and let i_0 be the smallest integer such that $(D+1)i_0 \geq t_0$. Then, we have that:

$$\sum_{i=0}^{+\infty}(1-\lambda(P(k_i,k_{i+1}))) \geq |\mathcal{B}| \sum_{i=i_0}^{+\infty} (\frac{\epsilon((D+1)i)}{5|\mathcal{C}|-1})^{(D+1)N}$$
$$= \frac{|\mathcal{B}|}{(5|\mathcal{C}|-1)^{(D+1)N}} \sum_{i=i_0}^{+\infty} \frac{1}{(D+1)i} = +\infty. \tag{4.3}$$

Hence, the weak ergodicity property follows from Theorem 6.4.□

All the conditions in Theorem 6.5 hold. Thus it follows from Theorem 6.5 that the limiting distribution is $\mu^* = \lim_{t\to+\infty} \mu^t$. Note that $\lim_{t\to+\infty} \mu^t = \lim_{t\to+\infty} \mu(\epsilon(t)) = \mu(0)$ and Proposition 4.3 shows that the support of $\mu(0)$ is contained in the set $\operatorname{diag}\mathcal{E}(\Gamma_{\rm cov})$. Hence, the support of μ^* is contained in the set $\operatorname{diag}\mathcal{E}(\Gamma_{\rm cov})$, implying that $\lim_{t\to+\infty} \mathbb{P}(z(t) \in \operatorname{diag}\mathcal{E}(\Gamma_{\rm cov})) = 1$. This completes the proof.

4.2. Convergence analysis of the DIACL **Algorithm.** First of all, we employ Theorem 6.6 to study the convergence properties of the associated DHACL algorithm. This is essential to analyze the DIACL algorithm.

To simplify notations, we will use $s_i(t-1) := s_i(\gamma_i^{(2)}(t)+1)$ in the remainder of this section. Observe that $z(t) := (s(t-1), s(t))$ in the DHACL algorithm constitutes a Markov chain $\{\mathcal{P}_t^\epsilon\}$ on the space $\mathcal{B}'$.

LEMMA 4.4. *The Markov chain $\{\mathcal{P}_t^\epsilon\}$ is a regular perturbation of $\{\mathcal{P}_t^0\}$.*

Proof. Pick any two states $z^1 := (s^0, s^1)$ and $z^2 := (s^1, s^2)$ with $z^1 \neq z^2$. We have that $P^\epsilon_{z^1 z^2} > 0$ if and only if there is some $i \in V$ such that $s^1_{-i} = s^2_{-i}$ and one of the following occurs: $s_i^2 \in \mathcal{F}_i(a_i^1) \setminus \{s_i^0, s_i^1\}$, $s_i^2 = s_i^1$ or $s_i^2 = s_i^0$. In particular, the following holds:

$$P^{\epsilon}_{z^1 z^2} = \begin{cases} \eta_1, & s_i^2 \in \mathcal{F}_i(a_i^1) \setminus \{s_i^0, s_i^1\}, \\ \eta_2, & s_i^2 = s_i^1, \\ \eta_3, & s_i^2 = s_i^0, \end{cases}$$

where

$$\eta_1 := \frac{\epsilon^{m_i}}{N|\mathcal{F}_i(a_i^1) \setminus \{s_i^0, s_i^1\}|}, \quad \eta_2 := \frac{1-\epsilon^{m_i}}{N(1+\epsilon^{\rho_i(s^0,s^1)})}, \quad \eta_3 := \frac{(1-\epsilon^{m_i}) \times \epsilon^{\rho_i(s^0,s^1)}}{N(1+\epsilon^{\rho_i(s^0,s^1)})}.$$

Observe that $0 < \lim_{\epsilon \to 0^+} \frac{\eta_1}{\epsilon^{m_i}} < +\infty$. Multiplying the numerator and denominator of η_2 by $\epsilon^{\Psi_i(s^1,s^0)-(u_i(s^1)-\Delta_i(s^1,s^0))}$, we obtain

$$\eta_2 = \frac{1-\epsilon^{m_i}}{N} \times \frac{\epsilon^{\Psi_i(s^0,s^1)-(u_i(s^1)-\Delta_i(s^1,s^0))}}{\eta_2'},$$

where $\eta_2' := \epsilon^{\Psi_i(s^0,s^1)-(u_i(s^1)-\Delta_i(s^1,s^0))} + \epsilon^{\Psi_i(s^0,s^1)-(u_i(s^0)-\Delta_i(s^0,s^1))}$. Use

$$\lim_{\epsilon \to 0^+} \epsilon^x = \begin{cases} 1, & x = 0, \\ 0, & x > 0, \end{cases}$$

and we have

$$\lim_{\epsilon \to 0^+} \frac{\eta_2}{\epsilon^{\Psi_i(s^0,s^1)-(u_i(s^1)-\Delta_i(s^1,s^0))}} = \begin{cases} \frac{1}{N}, & u_i(s^0) - \Delta_i(s^0,s^1) \neq u_i(s^1) - \Delta_i(s^1,s^0), \\ \frac{1}{2N}, & \text{otherwise}. \end{cases}$$

Similarly, it holds that

$$\lim_{\epsilon \to 0^+} \frac{\eta_3}{\epsilon^{\Psi_i(s^0,s^1)-(u_i(s^0)-\Delta_i(s^0,s^1))}} \in \{\frac{1}{2N}, \frac{1}{N}\}.$$

Hence, the resistance of the feasible transition $z^1 \to z^2$, with $z^1 \neq z^2$ and sensor i as the unilateral deviator, can be described as follows:

$$\chi(z^1 \to z^2) = \begin{cases} m_i, & s_i^2 \in \mathcal{F}_i(a^1) \setminus \{s_i^0, s_i^1\}, \\ \Psi_i(s^0,s^1) - (u_i(s^1) - \Delta_i(s^1,s^0)), & s_i^2 = s_i^1, \\ \Psi_i(s^0,s^1) - (u_i(s^0) - \Delta_i(s^0,s^1)), & s_i^2 = s_i^0. \end{cases}$$

Then (A3) in Section 6.2 holds. It is straightforward to verify that (A2) in Section 6.2 holds. We are now in a position to verify (A1). Since $\mathcal{G}_{\text{loc}}$ is undirected and connected, and multiple sensors can stay in the same position, then $\diamond a^0 = \mathcal{Q}^N$ for any $a^0 \in \mathcal{Q}$. Since sensor i can choose any camera control vector from $\mathcal{C}$ at each time, then $\diamond s^0 = \mathcal{A}$ for any $s^0 \in \mathcal{A}$. This implies that $\diamond z^0 = \mathcal{B}'$ for any $z^0 \in \mathcal{B}'$, and thus the Markov chain $\{\mathcal{P}_t^{\epsilon}\}$ is irreducible on the space $\mathcal{B}'$.

It is easy to see that any state in $\operatorname{diag}\mathcal{A}$ has period 1. Pick any $(s^0, s^1) \in \mathcal{B}' \setminus \operatorname{diag}\mathcal{A}$. Since $\mathcal{G}_{\text{loc}}$ is undirected, then $s_i^0 \in \mathcal{F}_i(a_i^1)$ if and only if $s_i^1 \in \mathcal{F}_i(a_i^0)$. Hence, the following two paths are both feasible:

$$(s^0, s^1) \to (s^1, s^0) \to (s^0, s^1)$$
$$(s^0, s^1) \to (s^1, s^1) \to (s^1, s^0) \to (s^0, s^1).$$

Hence, the period of the state (s^0, s^1) is 1. This proves aperiodicity of $\{\mathcal{P}_t^\epsilon\}$. Since $\{\mathcal{P}_t^\epsilon\}$ is irreducible and aperiodic, then (A1) holds. $\square$

A direct result of Lemma 4.4 is that for each $\epsilon > 0$, there exists a unique stationary distribution of $\{\mathcal{P}_t^\epsilon\}$, say $\mu(\epsilon)$. From the proof of Lemma 4.4, we can see that the resistance of an experiment is m_i if sensor i is the unilateral deviator. We now proceed to utilize Theorem 6.6 to characterize $\lim_{\epsilon\to 0^+} \mu(\epsilon)$.

PROPOSITION 4.5. *Consider the regular perturbed Markov process $\{\mathcal{P}_t^\epsilon\}$. Then $\lim_{\epsilon\to 0^+} \mu(\epsilon)$ exists and the limiting distribution $\mu(0)$ is a stationary distribution of $\{\mathcal{P}_t^0\}$. Furthermore, the stochastically stable states (i.e., the support of $\mu(0)$) are contained in the set* $\operatorname{diag} S^*$.

Proof. The unperturbed Markov chain corresponds to the DHACL Algorithm with $\epsilon = 0$. Hence, the recurrent communication classes of the unperturbed Markov chain are contained in the set $\operatorname{diag}\mathcal{A}$. We will construct resistance trees over vertices in the set $\operatorname{diag}\mathcal{A}$. Denote $T_{\min}$ by the minimum resistance tree. The remainder of the proof is divided into the following four claims.

CLAIM 8. $\chi((s^0, s^0) \Rightarrow (s^1, s^1)) = m_i + \Psi_i(s^1, s^0) - (u_i(s^1) - \Delta_i(s^1, s^0))$ *where $s^0 \neq s^1$ and the transition $s^0 \to s^1$ is feasible with sensor i as the unilateral deviator.*

Proof. One feasible path for $(s^0, s^0) \Rightarrow (s^1, s^1)$ is $\mathcal{L} := (s^0, s^0) \to (s^0, s^1) \to (s^1, s^1)$ where sensor i experiments in the first transition and does not experiment in the second one. The total resistance of the path $\mathcal{L}$ is $m_i + \Psi_i(s^1, s^0) - (u_i(s^1) - \Delta_i(s^1, s^0))$ which is at most $m_i + m^*$.

Denote by $\mathcal{L}'$ the path with minimum resistance among all the feasible paths for $(s^0, s^0) \Rightarrow (s^1, s^1)$. Assume that the first transition in $\mathcal{L}'$ is $(s^0, s^0) \to (s^0, s^2)$ where node j experiments and $s^2 \neq s^1$. Observe that the resistance of $(s^0, s^0) \to (s^0, s^2)$ is m_j. No matter whether j is equal to i or not, the path $\mathcal{L}'$ must include at least one more experiment to introduce s_i^1. Hence the total resistance of the path $\mathcal{L}'$ is at least $m_i + m_j$. Since $m_i + m_j > m_i + 2m^*$, then the path $\mathcal{L}'$ has a strictly larger resistance than the path $\mathcal{L}$. To avoid a contradiction, the path $\mathcal{L}'$ must start from the transition $(s^0, s^0) \to (s^0, s^1)$. Similarly, the sequent transition (which is also the last one) in the path $\mathcal{L}'$ must be $(s^0, s^1) \to (s^1, s^1)$ and thus $\mathcal{L}' = \mathcal{L}$. Hence, the resistance of the transition $(s^0, s^0) \Rightarrow (s^1, s^1)$ is the total resistance of the path $\mathcal{L}$; i.e., $m_i + \Psi_i(s^1, s^0) - (u_i(s^1) - \Delta_i(s^1, s^0))$.$\square$

CLAIM 9. *All the edges $((s, s), (s', s'))$ in $T_{\min}$ must consist of only one deviator; i.e., $s_i \neq s'_i$ and $s_{-i} = s'_{-i}$ for some $i \in V$.*

Proof. Assume that $(s, s) \Rightarrow (s', s')$ has at least two deviators. Suppose the path $\hat{\mathcal{L}}$ has the minimum resistance among all the paths from (s, s) to (s', s'). Then, $\ell \geq 2$ experiments are carried out along $\hat{\mathcal{L}}$. Denote i_k by the unilateral deviator in the k-th experiment $s^{k-1} \to s^k$ where $1 \leq k \leq \ell$, $s^0 = s$ and $s^\ell = s'$. Then the resistance of $\hat{\mathcal{L}}$ is at least $\sum_{k=1}^{\ell} m_{i_k}$; i.e., $\chi((s^0, s^0) \Rightarrow (s', s')) \geq \sum_{k=1}^{\ell} m_{i_k}$.

Let us consider the following path on $T_{\min}$:

$$\bar{\mathcal{L}} := (s^0, s^0) \Rightarrow (s^1, s^1) \Rightarrow \cdots \Rightarrow (s^\ell, s^\ell).$$

From Claim 1, we know that the total resistance of the path $\bar{\mathcal{L}}$ is at most $\sum_{k=1}^{\ell} m_{i_k} + \ell m^*$.

A new tree T' can be obtained by adding the edges of $\bar{\mathcal{L}}$ into $T_{\min}$ and removing the redundant edges. The removed resistance is *strictly* greater than $\sum_{k=1}^{\ell} m_{i_k} + 2(\ell - 1)m^*$ where $\sum_{k=1}^{\ell} m_{i_k}$ is the lower bound on the resistance on the edge from (s^0, s^0) to (s^ℓ, s^ℓ), and $2(\ell - 1)m^*$ is the strictly lower bound on the total resistances of leaving (s^k, s^k) for $k = 1, \cdots, \ell - 1$. The adding resistance is the total resistance of $\bar{\mathcal{L}}$ which is at most $\sum_{k=1}^{\ell} m_{i_k} + \ell m^*$. Since $\ell \geq 2$, we have that $2(\ell-1)m^* \geq \ell m^*$ and thus T' has a strictly lower resistance than $T_{\min}$. This contradicts the fact that $T_{\min}$ is a minimum resistance tree.□

CLAIM 10. *Given any edge $((s,s),(s',s'))$ in $T_{\min}$, denote by i the unilateral deviator between s and s'. Then the transition $s_i \to s_i'$ is feasible.*

Proof. Assume that the transition $s_i \to s_i'$ is infeasible. Suppose the path $\check{\mathcal{L}}$ has the minimum resistance among all the paths from (s,s) to (s',s'). Then, there are $\ell \geq 2$ experiments in $\check{\mathcal{L}}$. The remainder of the proof is similar to that of Claim 9.□

CLAIM 11. *Let h_v be the root of $T_{\min}$. Then, $h_v \in \operatorname{diag} S^*$.*

Proof. Assume that $h_v = (s^0, s^0) \notin \operatorname{diag} S^*$. Pick any $(s^*, s^*) \in \operatorname{diag} S^*$. By Claim 9 and 10, we have that there is a path from (s^*, s^*) to (s^0, s^0) in the tree $T_{\min}$ as follows:

$$\tilde{\mathcal{L}} := (s^\ell, s^\ell) \Rightarrow (s^{\ell-1}, s^{\ell-1}) \Rightarrow \cdots \Rightarrow (s^1, s^1) \Rightarrow (s^0, s^0)$$

for some $\ell \geq 1$. Here, $s^* = s^\ell$, there is only one deviator, say i_k, from s^k to s^{k-1}, and the transition $s^k \to s^{k-1}$ is feasible for $k = \ell, \ldots, 1$.

Since the transition $s^k \to s^{k+1}$ is also feasible for $k = 0, \ldots, \ell-1$, we obtain the reverse path $\tilde{\mathcal{L}}'$ of $\tilde{\mathcal{L}}$ as follows:

$$\tilde{\mathcal{L}}' := (s^0, s^0) \Rightarrow (s^1, s^1) \Rightarrow \cdots \Rightarrow (s^{\ell-1}, s^{\ell-1}) \Rightarrow (s^\ell, s^\ell).$$

By Claim 8, the total resistance of the path $\tilde{\mathcal{L}}$ is

$$\chi(\tilde{\mathcal{L}}) = \sum_{k=1}^{\ell} m_{i_k} + \sum_{k=1}^{\ell} \{\Psi_{i_k}(s^k, s^{k-1}) - (u_{i_k}(s^{k-1}) - \Delta_{i_k}(s^{k-1}, s^k))\},$$

and the total resistance of the path $\tilde{\mathcal{L}}'$ is

$$\chi(\tilde{\mathcal{L}}') = \sum_{k=1}^{\ell} m_{i_k} + \sum_{k=1}^{\ell} \Psi_{i_k}(s^{k-1}, s^k) - (u_{i_k}(s^k) - \Delta_{i_k}(s^k, s^{k-1})).$$

Denote $\Lambda_1' := (\mathcal{D}(a_{i_k}^k, r_{i_k}^k) \backslash \mathcal{D}(a_{i_{k-1}}^{k-1}, r_{i_{k-1}}^{k-1})) \cap \mathcal{Q}$ and $\Lambda_2' := (\mathcal{D}(a_{i_{k-1}}^{k-1}, r_{i_{k-1}}^{k-1}) \backslash \mathcal{D}(a_{i_k}^k, r_{i_k}^k)) \cap \mathcal{Q}$. Observe that

$$\begin{aligned}
&U_g(s^k) - U_g(s^{k-1}) \\
&= u_{i_k}(s^k) - u_{i_k}(s^{k-1}) - \sum_{q \in \Lambda_1'} W_q\left(\frac{n_q(s^{k-1})}{n_q(s^{k-1})} - \frac{n_q(s^{k-1})}{n_q(s^k)}\right) + \sum_{q \in \Lambda_2'} W_q\left(\frac{n_q(s^k)}{n_q(s^k)} - \frac{n_q(s^k)}{n_q(s^{k-1})}\right) \\
&= (u_{i_k}(s^k) - \Delta_{i_k}(s^k, s^{k-1})) - (u_{i_k}(s^{k-1}) - \Delta_{i_k}(s^{k-1}, s^k)).
\end{aligned}$$

We now construct a new tree T' with the root (s^*, s^*) by adding the edges of $\tilde{\mathcal{L}}'$ to the tree $T_{\min}$ and removing the redundant edges $\tilde{\mathcal{L}}$. Since $\Psi_{i_k}(s^{k-1}, s^k) = \Psi_{i_k}(s^k, s^{k-1})$, the difference in the total resistances across the trees $\chi(T')$ and $\chi(T_{\min})$ is given by

$$\begin{aligned}
&\chi(T') - \chi(T_{\min}) = \chi(\tilde{\mathcal{L}}') - \chi(\tilde{\mathcal{L}}) \\
&= \sum_{k=1}^{\ell} -(u_{i_k}(s^{k-1}) - \Delta_{i_k}(s^{k-1}, s^k)) - \sum_{k=1}^{\ell} -(u_{i_k}(s^k) - \Delta_{i_k}(s^k, s^{k-1})) \\
&= \sum_{k=1}^{\ell} (U_g(s^k) - U_g(s^{k-1})) = U_g(s^0) - U_g(s^*) < 0.
\end{aligned}$$

This contradicts that $T_{\min}$ is a minimum resistance tree.□

It follows from Claim 4 that the state $h_v \in \operatorname{diag} S^*$ has minimum stochastic potential. Then Proposition 4.5 is a direct result of Theorem 6.6. □

We are now ready to show Theorem 3.2.

Proof of Theorem 3.1:

CLAIM 12. *Condition (B2) in Theorem 6.5 holds.*

Proof. The proof is analogous to Claim 5.□

CLAIM 13. *Condition (B3) in Theorem 6.5 holds.*

Proof. Denote by $P^{\epsilon(t)}$ the transition matrix of $\{\mathcal{P}_t\}$. Consider the feasible transition $z^1 \to z^2$ with unilateral deviator i. The corresponding probability is given by

$$P^{\epsilon(t)}_{z^1 z^2} = \begin{cases} \eta_1, & s_i^2 \in \mathcal{F}_i(a_i^1) \setminus \{s_i^0, s_i^1\}, \\ \eta_2, & s_i^2 = s_i^1, \\ \eta_3, & s_i^2 = s_i^0, \end{cases}$$

where

$$\eta_1 := \frac{\epsilon(t)^{m_i}}{N|\mathcal{F}_i(a_i^1) \setminus \{s_i^0, s_i^1\}|}, \quad \eta_2 := \frac{1-\epsilon(t)^{m_i}}{N(1+\epsilon(t)^{\rho_i(s^0,s^1)})}, \quad \eta_3 := \frac{(1-\epsilon(t)^{m_i}) \times \epsilon(t)^{\rho_i(s^0,s^1)}}{N(1+\epsilon(t)^{\rho_i(s^0,s^1)})}.$$

The remainder is analogous to Claim 6 in Section **??**.□

CLAIM 14. *Condition (B1) in Theorem 6.5 holds.*

Proof. Observe that $|\mathcal{F}_i(a_i^1)| \leq 5|\mathcal{C}|$. Since $\epsilon(t)$ is strictly decreasing, there is $t_0 \geq 1$ such that t_0 is the first time when $1 - \epsilon(t)^{m_i} \geq \epsilon(t)^{m_i}$.

Observe that for all $t \geq 1$, it holds that

$$\eta_1 \geq \frac{\epsilon(t)^{m_i}}{N(5|\mathcal{C}|-1)} \geq \frac{\epsilon(t)^{m_i+m^*}}{N(5|\mathcal{C}|-1)}.$$

Denote $b := u_i(s^1) - \Delta_i(s^1, s^0)$ and $a := u_i(s^0) - \Delta_i(s^0, s^1)$. Then $\rho_i(s^0, s^1) = b - a$. Since $b - a \leq m^*$, then for $t \geq t_0$ it holds that

$$\begin{aligned} \eta_2 &= \frac{1-\epsilon(t)^{m_i}}{N(1+\epsilon(t)^{b-a})} = \frac{(1-\epsilon(t)^{m_i})\epsilon(t)^{\max\{a,b\}-b}}{N(\epsilon(t)^{\max\{a,b\}-b} + \epsilon(t)^{\max\{a,b\}-a})} \\ &\geq \frac{\epsilon(t)^{m_i}\epsilon(t)^{\max\{a,b\}-b}}{2N} \geq \frac{\epsilon(t)^{m_i+m^*}}{N(5|\mathcal{C}|-1)}. \end{aligned}$$

Similarly, for $t \geq t_0$, it holds that

$$\eta_3 = \frac{(1-\epsilon(t)^{m_i})\epsilon(t)^{\max\{a,b\}-a}}{N(\epsilon(t)^{\max\{a,b\}-b} + \epsilon(t)^{\max\{a,b\}-a})} \geq \frac{\epsilon(t)^{m_i+m^*}}{N(5|\mathcal{C}|-1)}.$$

Since $m_i \in (2m^*, Km^*]$ for all $i \in V$ and $Km^* > 1$, then for any feasible transition $z^1 \to z^2$ with $z^1 \neq z^2$, it holds that:

$$P^{\epsilon(t)}_{z^1 z^2} \geq \frac{\epsilon(t)^{(K+1)m^*}}{N(5|\mathcal{C}|-1)}$$

for all $t \geq t_0$. Furthermore, for all $t \geq t_0$ and all $z^1 \in \operatorname{diag} \mathcal{A}$, we have that:

$$P^{\epsilon(t)}_{z^1 z^1} = 1 - \frac{1}{N}\sum_{i=1}^{N} \epsilon(t)^{m_i} = \frac{1}{N}\sum_{i=1}^{N}(1-\epsilon(t)^{m_i}) \geq \frac{1}{N}\sum_{i=1}^{N}\epsilon(t)^{m_i} \geq \frac{\epsilon(t)^{(K+1)m^*}}{N(5|\mathcal{C}|-1)}.$$

Choose $k_i := (D+1)i$ and let i_0 be the smallest integer such that $(D+1)i_0 \geq t_0$. Similar to (4.3), we can derive the following property

$$\sum_{\ell=0}^{+\infty}(1-\lambda(P(k_\ell,k_{\ell+1}))) \geq \frac{|\mathcal{B}|}{(N(5|\mathcal{C}|-1))^{(D+1)(K+1)m^*}}\sum_{i=i_0}^{+\infty}\frac{1}{(D+1)i} = +\infty.$$

Hence, the weak ergodicity of $\{\mathcal{P}_t\}$ follows from Theorem 6.4.□

All the conditions in Theorem 6.5 hold. Thus it follows from Theorem 6.5 that the limiting distribution is $\mu^* = \lim_{t\to+\infty}\mu^t$. Note that $\lim_{t\to+\infty}\mu^t = \lim_{t\to+\infty}\mu(\epsilon(t)) = \mu(0)$ and Proposition 4.5 shows that the support of $\mu(0)$ is contained in the set $\operatorname{diag} S^*$. Hence, the support of μ^* is contained in the set $\operatorname{diag} S^*$, implying that $\lim_{t\to+\infty}\mathbb{P}(z(t)\in\operatorname{diag} S^*) = 1$. It completes the proof.

5. Conclusions. We have formulated a coverage optimization problem as a constrained potential game. We have proposed two payoff-based distributed learning algorithms for this coverage game and shown that these algorithms converge in probability to the set of constrained NEs and the set of global optima of certain coverage performance metric, respectively.

6. Appendix. For the sake of a self-contained exposition, we include here some background in Markov chains [12] and the Theory of Resistance Trees [28].

6.1. Background in Markov chains. A *discrete-time Markov chain* is a discrete-time stochastic process on a finite (or countable) state space and satisfies the Markov property (i.e., the future state depends on its present state, but not the past states). A discrete-time Markov chain is said to be *time-homogeneous* if the probability of going from one state to another is independent of the time when the step is taken. Otherwise, the Markov chain is said to be *time-inhomogeneous*.

Since time-inhomogeneous Markov chains include time-homogeneous ones as special cases, we will restrict our attention to the former in the remainder of this section. The evolution of a time-inhomogeneous Markov chain $\{\mathcal{P}_t\}$ can described by the transition matrix $P(t)$ which gives the probability of traversing from one state to another at each time t.

Consider a Markov chain $\{\mathcal{P}_t\}$ with time-dependent transition matrix $P(t)$ on a finite state space X. Denote by $P(m,n) := \prod_{t=m}^{n-1}P(t)$, $0\leq m<n$.

DEFINITION 6.1 (Strong ergodicity [12]). *The Markov chain $\{\mathcal{P}_t\}$ is strongly ergodic if there exists a stochastic vector μ^* such that for any distribution μ on X and any $m\in\mathbb{Z}_+$, it holds that $\lim_{k\to+\infty}\mu^T P(m,k) = (\mu^*)^T$.*

Strong ergodicity of $\{\mathcal{P}_t\}$ is equivalent to $\{\mathcal{P}_t\}$ being convergent in distribution and will be employed to characterize the long-run properties of our learning algorithm. The investigation of conditions under which strong ergodicity holds is aided by the introduction of the coefficient of ergodicity and weak ergodicity defined next.

DEFINITION 6.2 (Coefficient of ergodicity [12]). *For any $n\times n$ stochastic matrix P, its coefficient of ergodicity is defined as $\lambda(P) := 1-\min_{1\leq i,j\leq n}\sum_{k=1}^n\min(P_{ik},P_{jk})$.*

DEFINITION 6.3 (Weak ergodicity [12]). *The Markov chain $\{\mathcal{P}_t\}$ is weakly ergodic if $\forall x,y,z\in X$, $\forall m\in\mathbb{Z}_+$, it holds that $\lim_{k\to+\infty}(P_{xz}(m,k)-P_{yz}(m,k)) = 0$.*

Weak ergodicity merely implies that $\{\mathcal{P}_t\}$ asymptotically forgets its initial state, but does not guarantee convergence. For a time-homogeneous Markov chain, there is no distinction between weak ergodicity and strong ergodicity. The following theorem provides the sufficient and necessary condition for $\{\mathcal{P}_t\}$ to be weakly ergodic.

THEOREM 6.4 ([12]). *The Markov chain $\{\mathcal{P}_t\}$ is weakly ergodic if and only if there is a strictly increasing sequence of positive numbers k_i, $i \in \mathbb{Z}_+$ such that $\sum_{i=0}^{+\infty}(1-\lambda(P(k_i,k_{i+1}))=+\infty$.*

We are now ready to present the sufficient conditions for strong ergodicity of the Markov chain $\{\mathcal{P}_t\}$.

THEOREM 6.5 ([12]). *A Markov chain $\{\mathcal{P}_t\}$ is strongly ergodic if the following conditions hold:*

(B1) The Markov chain $\{\mathcal{P}_t\}$ is weakly ergodic.

(B2) For each t, there exists a stochastic vector μ^t on X such that μ^t is the left eigenvector of the transition matrix $P(t)$ with eigenvalue 1.

(B3) The eigenvectors μ^t in (B2) satisfy $\sum_{t=0}^{+\infty}\sum_{z\in X}|\mu_z^t-\mu_z^{t+1}|<+\infty$.

Moreover, if $\mu^*=\lim_{t\to+\infty}\mu^t$, then μ^* is the vector in Definition 6.1.

6.2. Background in the Theory of Resistance Trees. Let P^0 be the transition matrix of the time-homogeneous Markov chain $\{\mathcal{P}_t^0\}$ on a finite state space X. And let P^ϵ be the transition matrix of a *perturbed Markov chain*, say $\{\mathcal{P}_t^\epsilon\}$. With probability $1-\epsilon$, the process $\{\mathcal{P}_t^\epsilon\}$ evolves according to P^0, while with probability ϵ, the transitions do not follow P^0.

A family of stochastic processes $\{\mathcal{P}_t^\epsilon\}$ is called a *regular perturbation* of $\{\mathcal{P}_t^0\}$ if the following holds $\forall x,y\in X$: (A1) For some $\varsigma>0$, the Markov chain $\{\mathcal{P}_t^\epsilon\}$ is irreducible and aperiodic for all $\epsilon\in(0,\varsigma]$.

(A2) $\lim_{\epsilon\to 0^+}P_{xy}^\epsilon=P_{xy}^0$.

(A3) If $P_{xy}^\epsilon>0$ for some ϵ, then there exists a real number $\chi(x\to y)\geq 0$ such that $\lim_{\epsilon\to 0^+}P_{xy}^\epsilon/\epsilon^{\chi(x\to y)}\in(0,+\infty)$.

In (A3), $\chi(x\to y)$ is called the resistance of the transition from x to y.

Let $H_1,H_2,\cdots,H_J$ be the recurrent communication classes of the Markov chain $\{\mathcal{P}_t^0\}$. Note that within each class H_ℓ, there is a path of zero resistance from every state to every other. Given any two distinct recurrence classes H_ℓ and H_k, consider all paths which start from H_ℓ and end at H_k. Denote $\chi_{\ell k}$ by the least resistance among all such paths.

Now define a complete directed graph $\mathcal{G}$ where there is one vertex ℓ for each recurrent class H_ℓ, and the resistance on the edge (ℓ,k) is $\chi_{\ell k}$. An *ℓ-tree* on $\mathcal{G}$ is a spanning tree such that from every vertex $k\neq\ell$, there is a unique path from k to ℓ. Denote by $G(\ell)$ the set of all *ℓ-trees* on $\mathcal{G}$. The resistance of an *ℓ-tree* is the sum of the resistances of its edges. The *stochastic potential* of the recurrent class H_ℓ is the least resistance among all *ℓ-trees* in $G(\ell)$.

THEOREM 6.6 ([28]). *Let $\{\mathcal{P}_t^\epsilon\}$ be a regular perturbation of $\{\mathcal{P}_t^0\}$, and for each $\epsilon>0$, let $\mu(\epsilon)$ be the unique stationary distribution of $\{\mathcal{P}_t^\epsilon\}$. Then $\lim_{\epsilon\to 0^+}\mu(\epsilon)$ exists and the limiting distribution $\mu(0)$ is a stationary distribution of $\{\mathcal{P}_t^0\}$. The stochastically stable states (i.e., the support of $\mu(0)$) are precisely those states contained in the recurrence classes with minimum stochastic potential.*

REFERENCES

[1] I.F. Akyildiz, T. Melodia, and K. Chowdhury. Wireless multimedia sensor networks: a survey. *IEEE Wireless Communications Magazine*, 14(6):32–39, 2007.

[2] T. Alpcan, T. Basar, and S. Dey. A power control game based on outage probabilities for multicell wireless data networks. *IEEE Transactions on Wireless Communications*, 5(4):890–899, 2006.

[3] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. June 2008. Manuscript preprint. Electronically available at http://coordinationbook.info.

[4] K.Y. Chow, K.S. Lui, and E.Y. Lam. Maximizing angle coverage in visual sensor networks. In *Proc. of IEEE International Conference on Communications*, pages 3516–3521, June 2007.

[5] J. Cortés, S. Martínez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM: Control, Optimisation & Calculus of Variations*, 11:691–719, 2005.

[6] R. Cucchiara. Multimedia surveillance systems. In *Proc. of the third ACM international workshop on Video surveillance and sensor networks*, pages 3–10, 2005.

[7] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems.* New York: Springer Verlag, 1984.

[8] D. Fudenberg and J. Tirole. *Game theory.* The MIT press, 1991.

[9] A. Ganguli, J. Cortés, and F. Bullo. Visibility-based multi-agent deployment in orthogonal environments. In *American Control Conference*, pages 3426–3431, New York, July 2007.

[10] A. Ganguli, J. Cortés, and F. Bullo. Multirobot rendezvous with visibility sensors in nonconvex environments. *IEEE Transactions on Robotics*, 25(2):340–352, 2009.

[11] E. Hörster and R. Lienhart. On the optimal placement of multiple visual sensors. In *Proc. of the third ACM international workshop on Video surveillance and sensor networks*, pages 111–120, 2006.

[12] D. Isaacson and R. Madsen. *Markov chains.* Wiley, 1976.

[13] A. Kwok and S. Martínez. Deployment algorithms for a power-constrained mobile sensor network. *International Journal of Robust and Nonlinear Control*, 2009. In press.

[14] K. Laventall and J. Cortés. Coverage control by multi-robot networks with limited-range anisotropic sensory. *International Journal of Control*, 2009. To appear.

[15] J. R. Marden, G. Arslan, and J. S. Shamma. Connections between cooperative control and potential games. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 2008. submitted.

[16] J. R. Marden and A. Wierman. Distributed welfare games. *Operations Research*, 2008. submitted.

[17] J.R. Marden and J.S. Shamma. Revisiting log-linear learning : asynchrony, completeness and payoff-based omplementation. *Games and Economic Behavior*, 2008. submitted.

[18] J.R. Marden, H.P. Young, G. Arslan, and J.S. Shamma. Payoff based dynamics for multi-player weakly acyclic games. *SIAM Journal on Control and Optimization*, 2006. submitted.

[19] C.B. Margi, V. Petkov, K. Obraczka, and R. Manduchi. Characterizing energy consumption in a visual sensor network testbed. In *Proc. of 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities*, pages 332–339, March 2006.

[20] D. Monderer and L. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[21] R. A. Murphey. Target-based weapon target assignment problems. In P. M. Pardalos and L. S. Pitsoulis, editors, *Nonlinear Assignment Problems: Algorithms and Applications*, pages 39–53. Kluwer Academic Publishers, 1999.

[22] J. O'Rourke. *Art Gallery Theorems and Algorithms.* Oxford University Press, 1987.

[23] R.W. Rosenthal. A class of games possseeing pure strategy nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[24] T. Roughgarden. *Selfish routing and the price of anarchy.* MIT press, 2005.

[25] T. C. Shermer. Recent results in art galleries. *Proceedings of the IEEE*, 80(9):1384–1399, 1992.

[26] A. Tang and L. Andrew. Game theory for heterogeneous flow control. In *Proc. of 42nd Annual Conference on Information Sciences and Systems*, pages 52–56, December 2008.

[27] J. Urrutia. Art gallery and illumination problems. In J. R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 973–1027. North-Holland, 2000.

[28] H.P. Young. The evolution of conventions. *Econometrica*, 61:57–84, Juanary 1993.